Environmental and Ecological Statistics **7,** 177–189, 2000

Sensitivity of normal theory methods to model misspecification in the calculation of upper confidence limits on the risk function for continuous responses

S.J. BANGA, G.P. PATIL and C. TAILLIE

Center for Statistical Ecology and Environmental Statistics, Department of Statistics, The Pennsylvania State University, University Park, PA 16802

Received December 1998; Revised May 1999

Normal theory procedures for calculating upper confidence limits (UCL) on the risk function for continuous responses work well when the data come from a normal distribution. However, if the data come from an alternative distribution, the application of the normal theory procedures may lead serious over- or under-coverage depending upon the alternative distribution. In this paper we conduct simulation studies to investigate the sensitivity of three normal theory UCL procedures to departures from normality. Data from several gamma, reciprocal gamma, and lognormal distributions are considered. The normal theory procedures are applied to both the raw data and the log-transformed data.

Keywords: coverage probability, dose-response models, gamma distribution, lognormal distribution, overcoverage, reciprocal gamma distribution, undercoverage

1. Introduction

Data analysis for the benchmark dose (BMD) with continuous responses (Crump, 1995; Chen and Gaylor, 1992; Kimmel and Gaylor, 1988) has generally been based on the normal model (Chen and Gaylor, 1992; Kodell and West, 1993a; Kodell and West, 1993a; Banga, Patil and Taillie, 1999a). One-sided confidence limits are usually employed as conservative point estimates. An upper confidence limit (UCL) on the risk function is of particular interest. The normal theory is appealing in this context but applying it where the data deviate from normality could lead to erroneous conclusions in the form of excessively stringent (and expensive) regulatory standards or excessively loose (and unsafe) standards. The effect of skewness on coverage probability is of particular concern.

Prepared with partial support from the Office of Research and Development, United States Environmental Protection Agency, Washington, DC under a Cooperative Agreement Number R825385. The contents have not been subjected to Agency review and therefore do not necessarily reflect the views of the Agency and no official endorsement should be inferred.

1352-8505 © 2000 Kluwer Academic Publishers

In order to study this issue, we consider the following two scenarios for constructing UCLs on the risk function. First, a normal based analysis is performed when the data actually follow either a gamma, a reciprocal gamma, or a lognormal distribution. In the second scenario, a lognormal analysis is conducted when the data actually come from either a gamma or a reciprocal gamma distribution. Our simulation study reveals a consistent pattern of overcoverage when the data are generated from the gamma model and of undercoverage when the data are generated from the reciprocal gamma or (first scenario only) from the lognormal model. Depending upon the model parameters, the under- or over-coverage can be quite severe and become progressively worse as the the sample size increases. Likelihood-based UCL methods developed for the gamma and reciprocal gamma distributions (Banga, Patil and Taillie, 1999c) provide a basis for doing the data analysis under some alternative (nonnormal) model specifications.

2. Dose-response models for continuous responses

Consider a dose-response experiment in which the response Y follows a continuous response distribution, $F(y;\boldsymbol{\psi}^*)$, some of whose parameters may vary with dose level d. A dose-response model for continuous responses thus involves two components: (i) the parametric family of response distributions $F(y;\boldsymbol{\psi}^*)$, and (ii) a ''link'' function specifying how $\boldsymbol{\psi}^*$ changes with the dose level. This latter specification typically involves unknown parameters $\boldsymbol{\theta}$ for example, the components of $\boldsymbol{\theta}$ might be the coefficients in a linear model. The full set of unknown parameters, denoted by $\boldsymbol{\psi}$, consists of $\boldsymbol{\theta}$ and any components of $\boldsymbol{\psi}^*$ that are not determined by $\boldsymbol{\theta}$ and d. Since there is a well-defined mapping $(\boldsymbol{\psi}, d)\mapsto\boldsymbol{\psi}^*$, the response distribution is indexed by $(\boldsymbol{\psi}, d)$ and we write $F(y;\boldsymbol{\psi}, d)$.

Risk analysis is brought into the picture by supposing that there is an ''abnormal'' point τ for which the outcome $Y \leq \tau$ indicates an adverse response. The risk function is defined as

$$R(d) \equiv R(d;\boldsymbol{\psi}) = \Pr(Y \leq \tau | d).$$

Here, we have taken the direction of adversity to the left; there is a parallel theory when the direction of adversity is to the right.

If τ were known numerically, then the observed responses could be dichotomized and the risk function estimated using quantal methods (although there is the question of whether statistical efficiency could be improved by using the continuous response directly). However, in this paper we examine the case in which no prior numerical value is available for τ. Instead, τ is defined as a quantile of the control distribution so that τ is itself an unknown parameter. Specifically, the investigator fixes a risk level α, with $0<\alpha<1$, and the abnormal point $\tau = \tau_\alpha$ is then determined by the requirement that

$$\alpha = \Pr(Y \leq \tau | d = 0).$$

Observe that τ depends upon the unknown parameters $\boldsymbol{\psi}$. On the other hand, the background risk, $R(0;\boldsymbol{\psi})$, equals α and is therefore known.

Our interest is in estimating the risk function $R(d;\boldsymbol{\psi})$ and more specifically in obtaining a (pointwise) upper confidence curve for that risk function.

Suppose that there are g experimental doses, $0 \leq d_1 < d_2 < \cdots < d_g$, with n_i observations, $Y_{i1}, Y_{i2}, \ldots, Y_{in_i}$, at dose level d_i. With the usual assumption of independent responses, the likelihood function is

$$L(\boldsymbol{\psi}) = \prod_{i=1}^{g}\prod_{j=1}^{n_i} f(y_{ij};\boldsymbol{\psi},d_i),$$

where $f(y;\boldsymbol{\psi},d)$ is the density corresponding to the response distribution $F(y;\boldsymbol{\psi},d)$. Banga *et al*. (1999a) describe a computationally convenient form of the profile likelihood method (called the likelihood contour method) for obtaining asymptotic UCLs on the risk function. Fairly explicit solutions are available for normally distributed responses with a linear link function. This is described in the next section.

3. Homoscedastic normal theory model for continuous responses

Suppose that responses are normally distributed with a constant (but unknown) variance and with a mean whose dependence on the dose is described by a linear model. Using the notation of the previous section, let

$$Y_{ij} \sim N(\mu(d_i),\sigma^2), \quad i = 1,\ldots,g, \quad j = 1,\ldots,n_i$$

with

$$\mu(d) = \boldsymbol{x}'\boldsymbol{\theta},$$

where $\boldsymbol{x}(d)$ is a p-dimensional vector whose components are known and $\boldsymbol{\theta}$ is a p-dimensional vector whose components are unknown. The unknown parameters are $\boldsymbol{\psi} = (\sigma^2,\theta_0,\ldots,\theta_{p-1})$. The total sample size is

$$N = \sum_{i=1}^{g} n_i,$$

where g is the number of experimental dose groups and n_i is the sample size for the ith group.

Specific examples of link functions include the straight line model with

$$\mu(d) = \theta_0 + \theta_1 d,$$

and the quadratic model with

$$\mu(d) = \theta_0 + \theta_1 d + \theta_2 d^2.$$

The quadratic model (Kodell and West, 1993a; Kodell and West, 1993b) allows the risk function to be decreasing for small dose levels in case the chemical under study is beneficial for small exposures.

Let $Y(d)$ be a hypothetical response at an arbitrary dose level d. Under the normal model

$$Y(d) \sim N(\mu(d),\sigma^2) \quad \text{and} \quad \mu(d) = \boldsymbol{x}'\boldsymbol{\theta},$$

so that the risk function becomes

$$R(d) = \Pr(Y(d) \le \tau | d) = \Phi\left(\frac{\tau - \mu(d)}{\sigma}\right).$$

The abnormal τ point is defined by

$$\alpha = R(0) = \Phi\left(\frac{\tau - \mu(0)}{\sigma}\right),$$

where α is specified by the investigator. The total risk can then be written as

$$R(d) = \Phi\left(z_\alpha + \frac{\mu(0) - \mu(d)}{\sigma}\right) = \Phi\left(z_\alpha + \frac{\boldsymbol{a}'\boldsymbol{\theta}}{\sigma}\right),$$

where

$$z_\alpha = \boldsymbol{\Phi}^{-1}(\alpha) \quad \text{and} \quad \boldsymbol{a} = \boldsymbol{x}(0) - \boldsymbol{x}(d).$$

We now outline three procedures for calculating an asymptotic UCL for $R(d)$.

3.1 *The MLE approach: MLE*

A standard method for calculating an asymptotic UCL for a scalar parameter uses asymptotic normality of the maximum likelihood estimator, and the so-called δ method to obtain an approximate standard error. For the homoscedastic normal model described above, this method yields an asymptotic $100(1-\alpha_c)$ percent UCL on $R(d)$ given by

$$\Phi\left(z_\alpha + \hat{\varphi} + z_{1-\alpha_c}\sqrt{\omega^2 + \hat{\varphi}^2/2N}\right),$$

where

$$\hat{\varphi} = \frac{\boldsymbol{a}'\hat{\boldsymbol{\theta}}}{\hat{\sigma}},$$

with $\hat{\boldsymbol{\theta}}$ and $\hat{\sigma}$ being the MLE estimates of $\boldsymbol{\theta}$ and σ, respectively. Also, $\omega^2 = \boldsymbol{a}'(X'X)^{-1}\boldsymbol{a}$, and $z_{1-\alpha_c}$ is the $100(1-\alpha_c)$th percentile of the standard normal distribution.

3.2 *Likelihood ratio-based approach: LREL and LRAL*

This approach uses the asymptotic χ^2 distribution of the likelihood ratio and the duality between confidence intervals and hypothesis testing to construct UCLs. Banga *et al.* (1999) formulate this as a constrained optimization problem which they solve using the Lagrange multiplier method. The exact solution to the Lagrange equations yields an asymptotic $100(1-\alpha_c)$ percent UCL for $R(d)$ given by

$$\Phi\left(z_\alpha + \frac{2\hat{\varphi} + \gamma^*\omega^2}{2\sqrt{1 - \gamma^*\hat{\varphi}/2N}}\right),$$

where γ^* is the larger root of the function

$$G(\gamma) = \left(1 - \frac{\hat{\varphi}\gamma}{2N}\right)\log\left(1 - \frac{\hat{\varphi}\gamma}{2N}\right) - \left(1 + \frac{z^2_{1-\alpha_c}}{N}\right)\left(1 - \frac{\hat{\varphi}\gamma}{2N}\right) + \frac{\gamma^2\omega^2}{4N} + 1.$$

This is referred to as the LREL procedure.

When the Lagrange equations are replaced by the lowest order terms in their asymptotic expansions, the resulting equations can be solved to yield another (less accurate) asymptotic $100(1-\alpha_c)$ percent UCL for $R(d)$ given by

$$\Phi\left(z_\alpha + \frac{\hat{\varphi}\sqrt{\omega^2 + \hat{\varphi}^2/2N} + z_{1-\alpha_c}\omega^2}{\sqrt{\omega^2 + \hat{\varphi}^2/2N} - z_{1-\alpha_c}\hat{\varphi}/2N}\right).$$

This is referred to as the LRAL approach.

4. Sensitivity scenario 1 (untransformed data)

For this scenario, we generate data as either gamma, reciprocal gamma, or lognormal but conduct the risk analysis as though the response distribution was actually normal as described in the previous section. The simulation studies were conducted using Mathematica (Wolfram, 1996).

Study 1.1 *Gamma model misspecified as normal*

Six simulation studies are conducted, each of which is determined by the parameter k in the model

$$Y(d) \sim \text{ Gamma}(k, \lambda(d)) \quad \text{with} \quad E(Y(d)) = k\lambda(d) = e^{3-d-0.1d^2},$$

where the shape parameter k is fixed at 0.25, 0.5, 1.0, 1.5, 2.0 and 4.0. But, the investigator is unaware of these facts and models the response as

$$Y(d) \sim N(\mu(d), \sigma^2) \quad \text{with} \quad \mu(d) = E(Y(d)) = \beta_0 + \beta_1 d + \beta_2 d^2. \tag{1}$$

For each of the six models there are five dose groups: a control group ($d = 0$) and four treatment or experimental dose groups d_1, d_2, d_3, and $d_4 (d_1 < d_2 < d_3 < d_4)$. Each of the experimental dose levels is varied so that it yields the same true risk across the six experiments. With this set-up, experiment-to-experiment differences can be attributed to distributional changes as the parameter k changes in the study. We assume that small response values are adverse and that the background risk in each of the six experiments is specified as $\alpha = 0.05$. In addition, each dose group is simulated with first, 5 ($n_i = 5,\ N = 25$), then 10 ($n_i = 10,\ N = 50$), and finally 20 ($n_i = 20,\ N = 100$) observations. For reasons of space, coverage probabilities are reported in this paper only for sample sizes $N = 25$ and $N = 100$. Full results, including $N = 50$, are available in the technical report on which this paper is based (Banga *et al.*, 1999a). Asymptotic 95% UCLs in each of the 3 designs of the 6 studies are calculated for 4000 replicates using the misspecified normal model. The simulated coverage probability in each case is computed as the proportion of the 4000 simulations for which the resulting UCLs are greater or equal to the true risk (calculated from the true model). Since the targeted coverage probability is

95%, the simulation error for the computed coverage has a standard deviation given approximately as $\sqrt{0.95(0.05)/4000} = 0.003$ or 0.3% points. Coverage probabilities are plotted in Figs. 1 and 2.

Study 1.2 *Reciprocal gamma model misspecified as normal*

The same simulations as in Study 1.1 are carried out. However, the data is generated as reciprocal gamma, according to the model:

$$Y(d) \sim \text{ ReciprocalGamma}(k, \lambda(d)) \quad \text{with} \quad E(1/Y(d)) = k\lambda(d) = e^3 + d + 0.1d^2,$$

where the parameter k is fixed at 0.25, 0.5, 1.0, 1.5, 2.0 and 4.0. The direction of adversity for $Y(d)$ is to left and the background risk is $\alpha = 0.05$. Notice that this is the same as a direction of adversity to the right for the corresponding gamma variates $1/Y(d)$. Coverage probabilities are plotted in Figs. 3 and 4 when the UCLs are calculated using the normal theory methods.

Study 1.3 *Lognormal misspecified as normal*

The true model here is

$$Y(d) \sim \text{ Lognormal}(\mu(d), \sigma^2) \quad \text{with} \quad E(\log Y(d)) = \mu(d) = 3 - d - 0.1d^2,$$

where the parameter σ is fixed at 0.25, 0.5, 1.0, 1.5, 2.0 and 4.0.

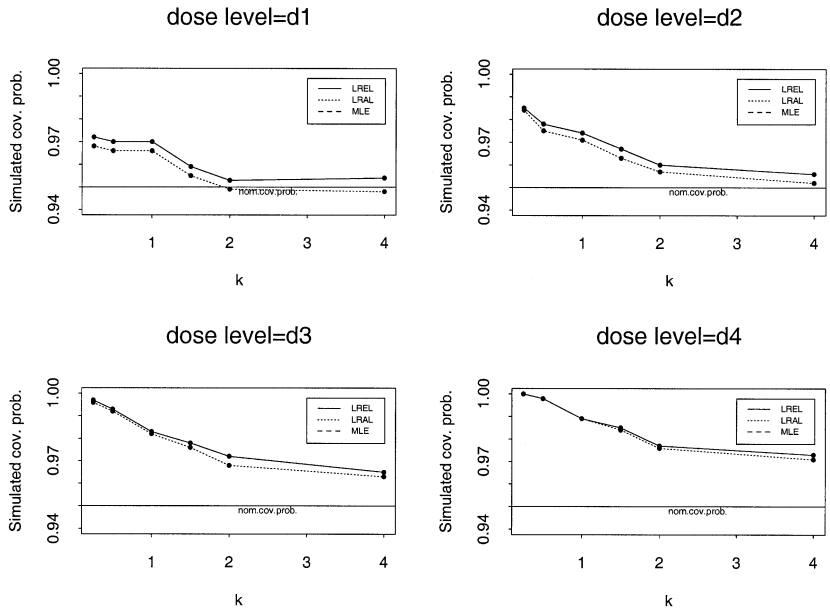

Figure 1. Study 1.1—The normal theory procedures are applied to the gamma models and the simulated coverage probabilities are plotted against the shape parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **5** animals so that the total sample size is $N = \mathbf{25}$.

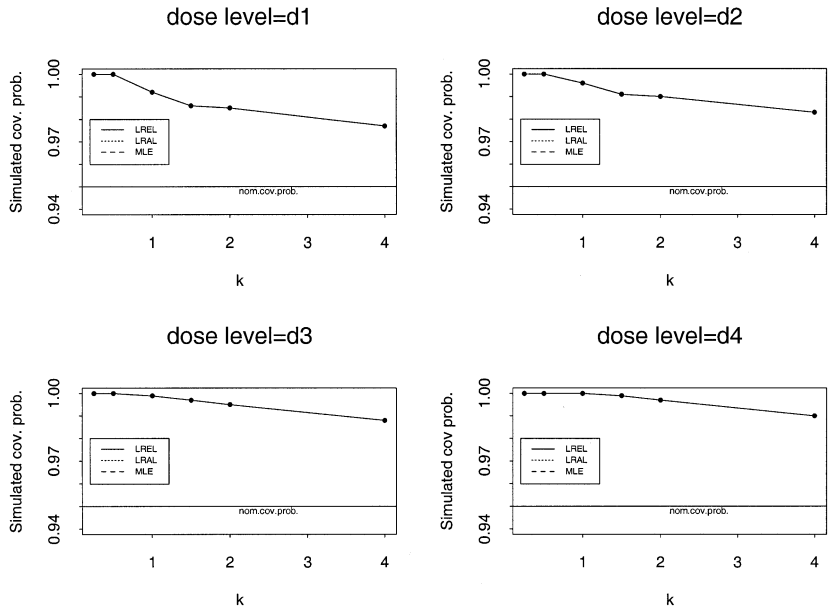

Figure 2. Study 1.1—The normal theory procedures are applied to the gamma models and the simulated coverage probabilities are plotted against the shape parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **20** animals so that the total sample size is $N = \mathbf{100}$.

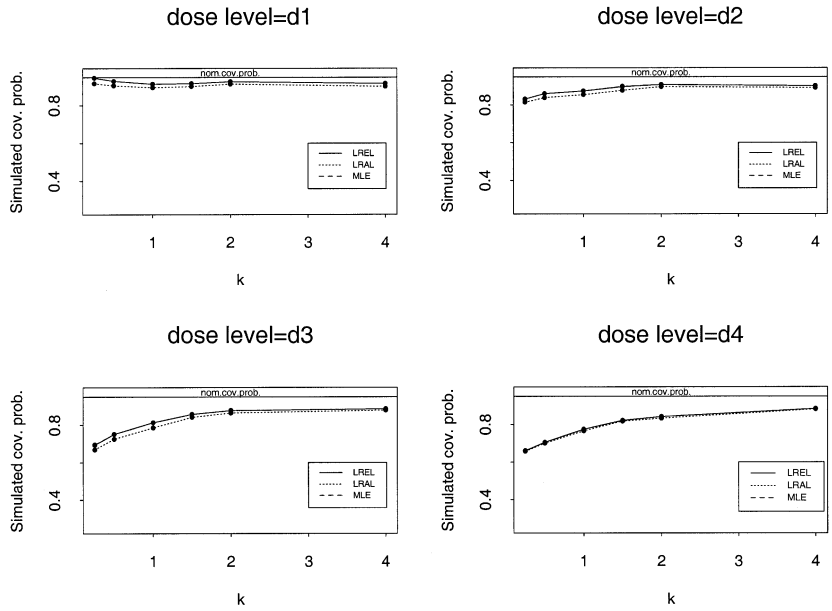

Figure 3. Study 1.2—Simulated coverage probabilities generated by the application of the normal theory methods to the reciprocal gamma models plotted against the model parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **5** animals so that the total sample size is $N = \mathbf{25}$.

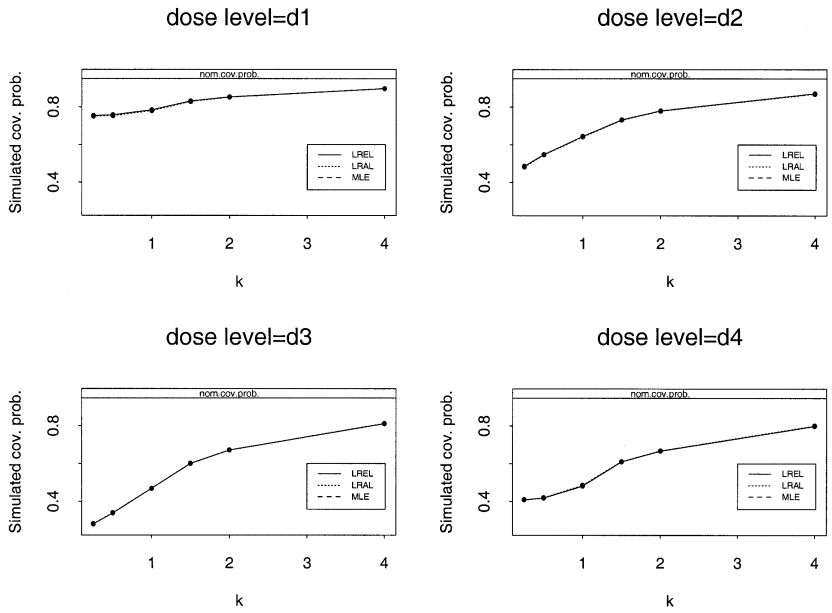

Figure 4. Study 1.2—Simulated coverage probabilities generated by the application of the normal theory methods to the reciprocal gamma models plotted against the model parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **20** animals so that the total sample size is $N = \mathbf{100}$.

Once again the direction of adversity is to the left and the background risk is $\alpha = 0.05$. Coverage probabilities are plotted in Figs. 5 and 6 when UCLs are calculated using the normal theory methods.

5. Sensitivity scenario 2 (log transformed data)

In this scenario, we generate the data from either the gamma or reciprocal gamma distribution but we analyze the data as if it were lognormal. Equivalently, apply a normal based analysis to the logarithmically transformed data. The simulations are conducted in the same set-up as in the first scenario described earlier. Again we examine the effect of such a misspecification on the simulated coverage probabilities of the UCLs. These are plotted in Figs. 7 and 8 for the gamma model (Study 2.1) and in Figs. 9 and 10 for the reciprocal gamma model (Study 2.2).

6. Results and discussion

The simulation results show that the achieved coverage probabilities for the three normal theory methods are extremely sensitive to departures from normality. The MLE and the LREL methods yield essentially the same UCLs and coverage probabilities. The LRAL method, on the other hand, provides UCLs and coverage probabilities that are slightly

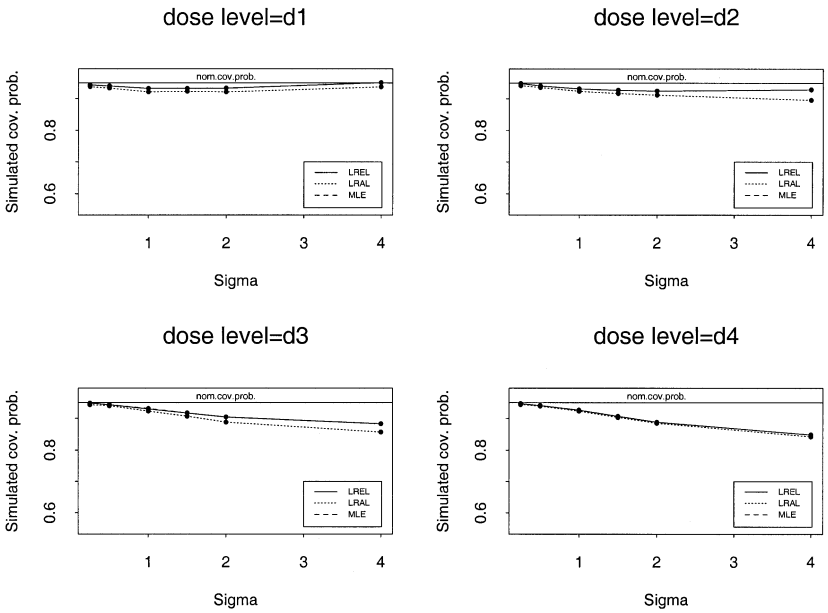

Figure 5. Study 1.3—The normal theory procedures are applied to the lognormal models and the simulated coverage probabilities are plotted against the parameter σ of the lognormal model. The targeted coverage probability is 0.95. Each experimental dose group consists of **5** animals so that the total sample size is $N = \mathbf{25}$.

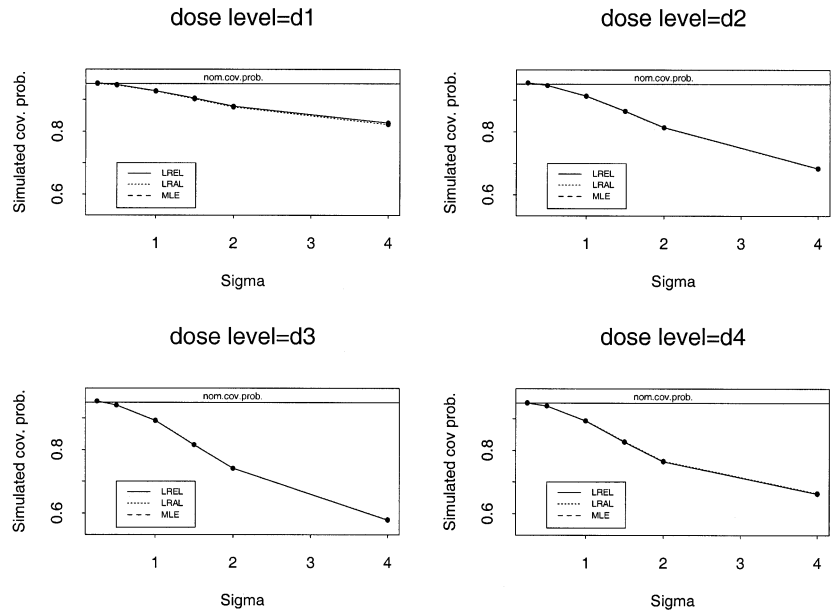

Figure 6. Study 1.3—The normal theory procedures are applied to the lognormal models and the simulated coverage probabilities are plotted against the parameter σ of the lognormal model. The targeted coverage probability is 0.95. Each experimental dose group consists of **20** animals so that the total sample size is $N = \mathbf{100}$.

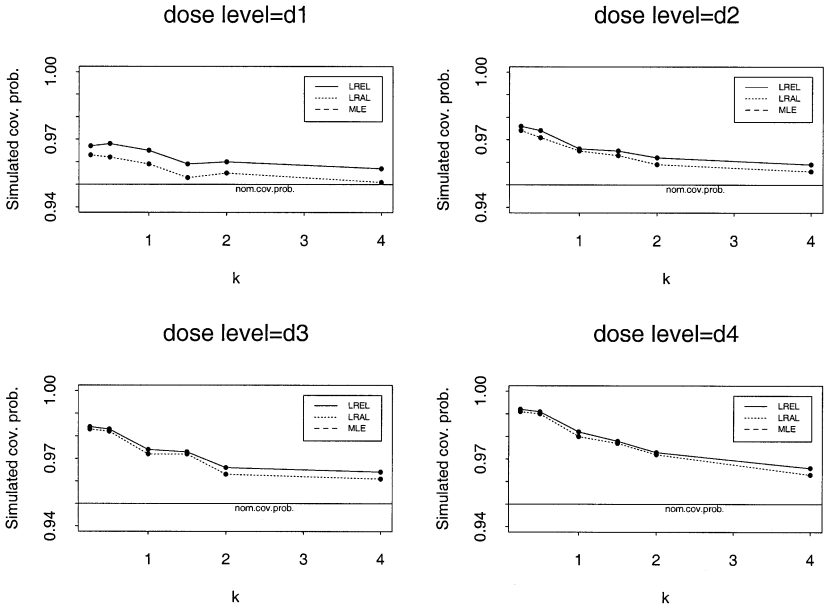

Figure 7. Study 2.1—The lognormal theory procedures are applied to the gamma models and the simulated coverage probabilities are plotted against the model parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **5** animals so that the total sample size is $N = \mathbf{25}$.

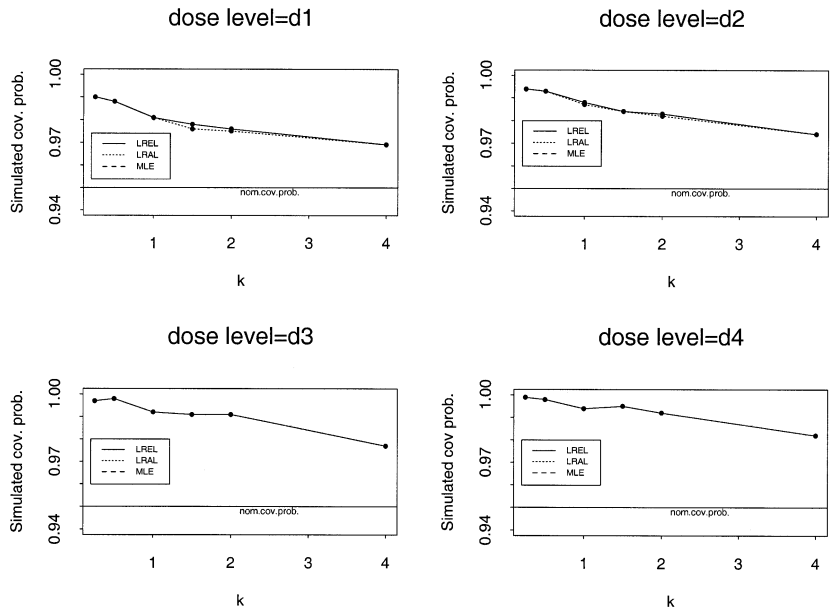

Figure 8. Study 2.1—The lognormal theory procedures are applied to the gamma models and the simulated coverage probabilities are plotted against the parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **20** animals so that the total sample size is $N = \mathbf{100}$.

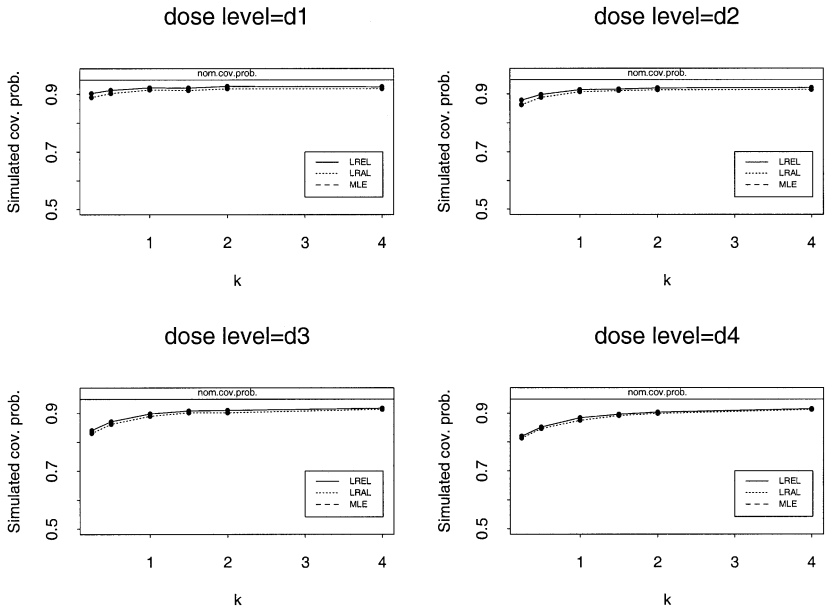

Figure 9. Study 2.2—Simulated coverage probabilities generated by the application of the lognormal theory methods to the reciprocal gamma models plotted against the parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **5** animals so that the total sample size is $N = \mathbf{25}$.

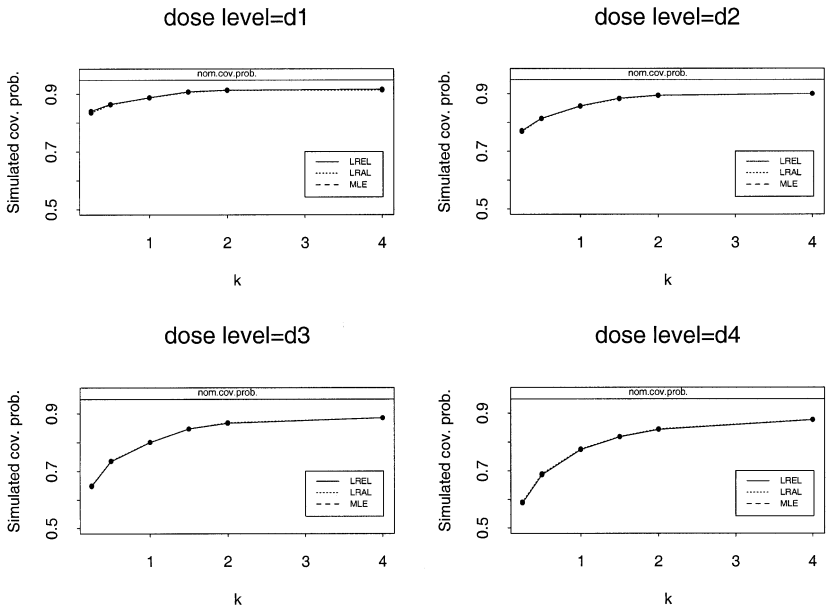

Figure 10. Study 2.2—Simulated coverage probabilities generated by the application of the lognormal theory methods to the reciprocal gamma models plotted against the parameter k. The targeted coverage probability is 0.95. Each experimental dose group consists of **20** animals so that the total sample size is $N = \mathbf{100}$.

different for small sample sizes but which quickly converge to those provided by the MLE and LREL methods as the sample size increases.

In Study 1.1, where the normal theory procedures are applied to the gamma models, the simulated coverage probabilities are unusually conservative. Further, the overcoverage becomes more pronounced as the sample size increases (see Figs. 1 and 2). We speculate that this is due to inconsistency of the normal theory estimators when applied to nonnormal data. The coverage probabilities are frequently above 0.99 for larger sample sizes. This effect is somewhat mitigated in Study 2.1 where the same procedures are applied to the logarithmic transformed gamma data (see Figs. 7 and 8). However, the resulting coverage probabilities are still highly conservative. In both of these studies, the coverage probabilities approach the nominal level as the shape parameter k increases. This is not surprising since the gamma distribution approaches the normal distribution as $k \to \infty$. The performance of the procedures in Study 1.2 are also adversely affected by the model misspecification. However, here the simulated coverage probabilities are liberal. These probabilities fall far below the target coverage as the sample size increases (see Figs. 3 and 4). This pattern is also observed in Study 2.2 where the same normal theory procedures are applied to the logarithmic transformed reciprocal data (see Figs. 9 and 10). The procedures perform slightly better in comparison to Study 1.2, but the simulated coverage probabilities are still very liberal. In both of these studies (Study 1.2 and Study 2.2), the simulated coverage probabilities approach the nominal value from below as the shape parameter k increases.

The simulated coverage probabilities are also excessively liberal when the normal methods are applied to the lognormal model. Figs. 5 and 6 suggest that for small values of σ (small skewness values) the simulated coverage probabilities are very close to the nominal level but then fall far below it as σ increases.

Interestingly, in all five studies the simulated coverage probabilities of the UCLs at lower dose levels seem to be only mildly affected by the model misspecification. The effect of the departures from normality on the simulated coverage probabilities becomes increasingly severe for higher dose levels.

Finally, we note that the normal theory methods result in overcoverage for the (short-tailed) gamma distribution, but in undercoverage for the (long-tailed) reciprocal gamma and lognormal distributions. It would be interesting to characterize those distributions which give overcoverage versus those which give undercoverage.

The preceding conclusions are also supported by the sample size $N = 50$, as reported in Banga *et al*. (1999b).

References

Banga, S.J., Patil, G.P., and Taillie, C. (1999) *Continuous Dose Response Modeling and Methodology for Nonnormal Distributions with Emphasis on Gamma and Reciprocal Gamma Distributions*, Technical Report 99–0402, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Penn State University, University Park, PA.

Banga, S.J., Patil, G.P., and Taillie, C. (1999) *Likelihood Contour and Lagrange Multiplier Method for the Calculation of Asymptotic Upper Confidence Limits on Risk Function for Quantitative Responses*, Risk Analysis, Under Revision.

Banga, S.J., Patil, G.P., and Taillie, C. (1999) *Sensitivity of Normal Theory Methods to Model Misspecification in the Calculation of Upper Confidence Limits on the Risk Function for Continuous Responses*, Technical Report 99–0403, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Penn State University, University Park, PA.

Chen, J.J. and Gaylor, D.W. (1992) Dose Response Modeling of Quantitative Response Data for Risk Assessment. *Communications in Statistics—Theory and Methods*, **21**(8), 2367–81.

Crump, K.S. (1995) Calculation of Benchmark doses from Continuous Data. *Risk Analysis*, **15**, 79–89.

Kimmel, C.A. and Gaylor, D.W. (1988) Issues in Qualitative and Quantitative Risk Analysis for Developmental Toxicity. *Risk Analysis*, **8**, 15–20.

Kodell, R.L. and West, R.W. (1993) Statistical Methods of Risk Assessment for Continuous Variables. *Communications in Statistics—Theory and Methods*, **22**(12), 3363–76.

Kodell, R.L. and West, R.W. (1993) Upper Confidence Limits on Excess Risk for Quantitative Responses. *Risk Analysis*, **13**(2), 177–82.

Wolfram, S. (1996) *The Mathematica Book,* 3rd edition, Wolfram Media/Cambridge University Press.

Biographical sketches

S.J. Banga is a Graduate Student in the Department of Statistics at the Pennsylvania State University.

G.P. Patil is Distinguished Professor of Mathematical Statistics and Director of the Center for Statistical Ecology and Environmental Statistics in the Department of Statistics at the Pennsylvania State University.

C. Taillie is Senior Research Associate in the Center for Statistical Ecology and Environmental Statistics in the Department of Statistics at the Pennsylvania State University.